10

Problem Solving as a Continuous Principle for Teaching

Suggestions and Examples

HANS HUMENBERGER

HANS-CHRISTIAN REICHEL

1. Introduction

It is not the aim of this chapter to discuss problem solving as generally as possible or to deal with strategies of problem solving in a systematic way. Many authors already have dealt with all these topics in various ways (e.g., Dürschlag, 1983; Engel, 1979; Polya, 1954, 1980; Sell, 1988; Zimmermann, 1983).

The emphasis of this contribution is clearly on a more or less systematic presentation of selected examples (problems), including detailed discussions of their solutions. We would consider ourselves most successful if teachers were to make use of some of these problems in their classes or if they were stimulated to look for, or to create, similar problems themselves so that students can profit from a lot of (we hope positive) experience in solving such problems. In that way, we

would make a small contribution to mathematics as well as to the teaching of mathematics.

We begin with some general introductory notes. Problem solving is an old topic in the teaching of mathematics, but it is not dealt with in a completely satisfying way. Many mathematicians and mathematics teachers dedicated their work to "problem solving" or to "drawing plausible conclusions," respectively. Many "strategies" among a large number of possibilities dealing with problem solving were (and are) emphasized as especially important; they are often demonstrated by some tasks, and sometimes students even get a "catalogue" on how to proceed and are then expected to find systematically the adequate strategies for solving the problems. Nevertheless, the following questions seem to us to remain not yet fully answered: Is it possible to teach heuristcs or problem solving systematically and to integrate it into the curriculum? Can a problem-solving process really be divided into "phases"? Is it possible to ensure a successful solution of a problem by a "strategy-pattern"?

We don't propose to give a definitive answer to these questions! We do not know if the "phases" of problem solving manifest themselves for everybody in the same way, if they appear in an ordered and comprehensible way at all, or if some phases appear parallel, disordered, and indescribable. A satisfying comprehension of a problem-solving process is hardly possible, because our thinking often appears in an un- or subconscious way; it is often not possible to say how one got certain ideas, or why one chose a certain way. In many cases one is working unfounded, trying and choosing ways without knowing where they will lead.

Despite all uncertainties about the systematics of problem solving, we do have the conviction that it is very important to provide enough opportunities in mathematics classes to solve problems that do not always fit into learned patterns, to be creative, to give reasons (to argue), to try (possibly wrong) ways, to get experiences, and so on. It seems a common practice among teachers to restrict problem solving to the better students while presenting students in the average compulsory classes with certain mathematical methods and later drilling them. We are not against "drilling certain methods"—quite the contrary—but we feel this should not be all of the classwork.

It is very important to find tasks that are "problems" in the sense that they promote creativity and spirit, that they motivate, that they are not limited to simple exercises but also are not too tricky (thus giving average students a realistic chance). Students can get self-confidence and be further motivated by having successful experiences. We are thinking of motivating and interesting tasks that contain "real mathematics" that fit into the curriculum, and that give students an appetite for more mathematics—not only gifted students but also for average and below average students!

We would not go so far as to suggest that all classwork should be based on problem solving, but we believe that one can (should) pose suitable and interesting problems for each mathematical topic, as "application tasks" or, later, as "repetition tasks," which show that the acquired mathematical knowledge was not just

important "for the test." This should start with the beginning of mathematics classes and should recur as a continuous class principle. Furthermore, it seems important to us that the students get enough time to deal with the problems themselves and that various possibilities are discussed thoroughly and are really understood by most of the students. There must be time for mistrials as well, and these must not simply be considered as "wrong"; there has to be an explanation why certain methods do not lead further—the problem must really be "digested"! It is of only little use for students to be rushed from one problem to another, or to have the teacher distinguish himself or herself with brilliantly demonstrated solutions (which appear out of nowhere) and where the student cannot even guess where the statement came from or why the teacher tried one way and not another. Learning to solve problems surely does not depend on which problems the students have to solve, as much as on how the lessons are shaped, what activities are expected from the students, and what the climate is in the classroom.

But the treatment of problem solving as a "leading idea" for classwork is not without difficulties, and we hope our chapter and this book will alleviate some of these problems:

- Tasks (problems) with a suitable degree of difficulty that are interesting and motivating are often not sufficiently presented in textbooks, and it is often not easy for teachers to find or create such tasks. Therefore, this chapter will offer stimulation in the form of a variety of mostly annotated examples that are aimed not at performance-oriented math team groups but primarily at "normal" mathematics lessons in schools.
- The judgment of whether or not a task is "suitable," in the above-mentioned way, is naturally a subjective matter, so that there does not have to be agreement between the authors or editors of particular collections of examples, teachers, and students who may consider some suggested problems boring or too hard or not usable for other reasons.
- As we mentioned already, a lot of problems can be solved only with the help of some tricks, but to draw the line between a trick and an important mathematical method is not easy and is also very subjective.

To summarize, we want to stress once more that it is definitively not enough to demonstrate problem solving once or twice a year to the students or to show them how a teacher is easily able to solve even difficult problems. Students must have various opportunities to get down to easier problems themselves so that they get a chance to think about them and to develop solutions themselves (maybe with some minor hints). Even if these solutions are not always correct, students can develop confidence in their own abilities and lead them to find the courage and joy to do more problems or more mathematics in general.

The following tasks are not ordered by diverse heuristic strategies but by two different aspects. The first section will deal with problems in which a colloquial text has to be translated into the "language of mathematics" (in most cases this will mean an equation, but not always!). In our opinion, this type of exercise should

be practiced particularly early and often. The translation of a colloquial problem into the language of mathematics is frequently "half of the solution" to a problem (cf. Reichel, 1991).

The second section consists of examples that clearly illustrate certain theories (or parts of theories) and help to make them more understandable. Moreover, they could promote motivation with their surprising results, and they might be well suited as introductory problems in certain fields. Even new mathematical topics can be introduced by simple problems, which show clearly the necessity and possibility of an extension of students' current knowledge with a generalization or a new theoretical concept. But it is exactly this type of problem that is often hard to find.

2. Translation of Texts Into the Language of Mathematics

One very important type of problem solving is translating colloquially described situations ("texts") into the language of mathematics. In such problems, it is first necessary to work out the mathematical structure of the task. One has to get an overview of the described situation, although it may not be immediately clear to show how it is possible to apply mathematics. It is especially important to deal with problems that allow more than one approach to the solution as well as problems that might contain superfluous data for students to detect. However, they should not be so complex that only extremely gifted students can manage them. The mathematical analysis of a problem will, in many cases, lead to an equation, but not necessarily (e.g., Example 2.3).

One wrong idea we want to confront decisively is that problem solving is interesting only for those students who have already achieved a relatively large, basic knowledge and who are at least 14 or 15 years old. It is also possible to formulate interesting problems for younger students (for instance, 10-year-olds) on a lower mathematical level. Simpler problems of this kind can be motivating for these students and can even help them in their further mathematical education. In the following sections, we want to present some elementary examples that can be used for students as young as 10.

2.1. *Simple Overdetermined Problems*

Working with over- and underdetermined problems is, in our opinion, a fine means to "sharpen the spirit," to promote the ability of the students to better recognize the structure (the essential things, the gist), and to translate (colloquially formulated) texts into questions that can be answered in a mathematical way. This capability is combined with the ability to apply mathematics—an ability everybody will need to solve his or her (mathematical) problems (in school and later, too) and which we should therefore master!

Example 2.1

ARITH

14

The freshman class of a high school in a small town of 4,300 inhabitants takes a trip to a mountain 120 km away. There is $500 in the class treasury. The total cost of this trip was $360. This amount was needed to pay the bus fee of $110 and to pay the cost of the rope-walk for each of the 25 students.

a. How much was the rope-walk for one student?
b. Which data were not necessary to answer question *a?*
   Rewrite the text using variables instead of specific numbers (a, b, c, . . .) and give a general formula for the required fee.

Here the students first have to find out which data are not necessary (freshman class, 4,300 inhabitants, 120 km, $500). Then they have to recognize the mathematical structure of the task: The total spending consists of the fees for the bus and rope-walk. The students should be able to find the formula for the student's fee p basically by themselves: $p = (360 - 110)/25 = 10$.

Such examples can be embellished and extended almost limitlessly, and there are no bounds for the teacher's imagination. The more sophisticated and complex the story is and the more details (variables or concrete numbers) that appear in the text, the more difficult it is, in general, to find the unnecessary data, particularly if several quantities are asked for. (Many students can visualize a situation much better with concrete numbers.) The students could also be asked to develop such "task-stories" themselves (e.g., in smaller groups where students then solve each other's tasks or as homework.) This may be a small contribution to promoting the student's motivation, especially the very young student's (10 years old).

Analogous to this, tasks with too little information for the solution—"underdetermined" problems—would be valuable too. Here the students would have to figure out why the problem is not solvable, and they could make suggestions as to what additional information is needed to solve the problem.

Reality—in contrast to math textbooks—often does not provide the full amount of data needed to solve given questions. On the other hand, we often have much more information than necessary. Therefore, to educate in the spirit of applied mathematics, textbooks should copy reality in this respect (even artificially) so that students have to figure out which and how many data have to be known to answer the questions asked before they start working. Students should also be able to say explicitly which additional data would be necessary. Such problems are called over- and underdetermined problems in this chapter.

Example 2.2

ARITH

An employee is going to work by bicycle. Usually he goes the 3 km distance at an average speed of 15 km/h. But this time he was unlucky because, after 1 kilometer, he got a flat tire, so the journey took him 20 minutes more because he had to push his bike from this point. Fortunately, he was able to repair the damage at work and could cycle home as usual.

a. How many kilometers more did he go by bike than he walked?
b. Which data are irrelevant to answer question *a?*
c. Which additional questions could be asked and answered?
   or
d. Is it possible to answer the following questions with the given data?
   (i) How long does the bike trip normally take him?
   (ii) How long (how much time) did he have to push his bike?
   (iii) What was the average walking speed?
   (iv) How long did the repair take him?

Solution: A sketch showing the distances that were walked and cycled may be very helpful (Figure 10.1).

6

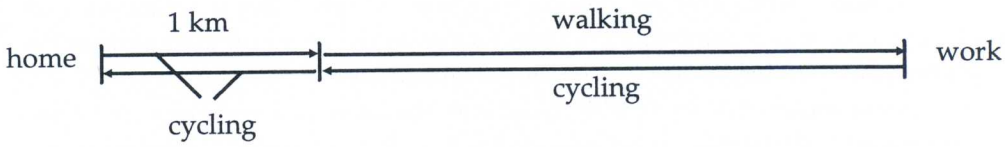

Figure 10.1 A cyclist who had some bad luck on his journey

To solve *(a)*: He had to push his bike from the scene of the accident to his work, but on his way back, he rode his bike exactly the same distance as he had to walk earlier (therefore the distance does not matter). The distance from home to the scene of the flat tire is covered twice by bike, once on the way out and once on the way back, so he went 2×1 km = 2 km more by bike than he walked.

14

To solve *(b)*: Everything is unnecessary except the 1 km! The value in question *(a)* does not even depend on the total distance of the trip.

To answer the four parts of *(d)*:

(i) $\frac{3}{15}h = 12$ min

(ii) 28 min

(iii) $4\frac{2}{7}$ km/h

(iv) not answerable

Example 2.3

Mr. Mayer usually takes the 7:25 bus from home to his office, and in the afternoon he walks home (for health reasons). This takes him altogether 1 hour and 10 minutes. If he were walking both to and from the office, it would take him 1 hour and 50 minutes, because on average he walks only 15 km/h slower than the bus travels, with its scheduled stops, traffic lights, congestion, and so forth. ARITH

a. How long would it take him to go both directions by bus?
b. What data are unnecessary to answer question *(a)?*
c. What questions could be answered with the given information?
   or
d. Is it possible to answer the following questions with the given information?
   (i) When does Mr. Mayer arrive at his office?
   (ii) How fast does Mr. Mayer walk, on average?
   (iii) How fast does the bus travel, on average?
   (iv) What is the distance from Mr. Mayer's office to his home?

Solution: For *(a)*; As walking both ways takes him altogether 1 hour 50 min, walking one way takes him 110/2 = 55 min. Therefore, one way takes the bus 70 min – 55 min = 15 min. Hence, both directions would take the bus 30 min.

To solve *(b)*: The data 7:25 and 15 km/h are superfluous.
To answer the four parts of *(d)*:

(i) Normally he arrives at his office at 7:40.

(ii) $5\frac{5}{8}$ km/h

(iii) $20\frac{5}{8}$ km/h

(iv) $5\frac{5}{32}$ km.

The latter three solutions are to be expected at the earliest from 12- or 13-year-olds.

2.2 *Problem Solving by Equations—Text Equations*

ALG

The following example will show, among other things, that it is possible, even at an elementary level, to demonstrate how important it is to mark the right quantities with variables.

Example 2.4

Forty-two birds are sitting on three trees. If 3 birds fly from the first tree to the second one and 7 birds fly from the second tree to the third one, there are twice as many birds on the second tree as on the first one and twice as many on the third tree as on second one. How many birds were sitting originally on each tree?

Probably a student's most familiar method of translating texts into the language of mathematics—to define the original bird-numbers with x, y, and z—will lead to a system of three linear equations with three unknowns:

$$x + y + z = 42$$

$$2(x - 3) = y - 4$$

$$2(y - 4) = z + 7.$$

Their solutions would be $x = 9$, $y = 16$, and $z = 17$. But such equation systems are solvable only for students who are at least 14 or 15 years old (at least in Austria). So one could ask the students to look for a solution that does not involve solving a system of equations. Alternatively, if this problem is stated already for 12- or 13-year-old students, then they are forced to find a different method anyway.

The situation, arithmetically, gets much easier if one looks at it first after the birds' flights: If there were a birds on the first tree, there were $2a$ birds on the second tree and $2\,(2a) = 4a$ birds on the third tree. That makes altogether $7a$ birds, whose total number is still 42. Therefore a—the number of birds on the first tree after the birds' flight—must be 6 ($42/7 = 6$). So after the birds' flight, there were 6, 12, and

1

24 birds on the trees 1, 2, and 3, respectively. Now we have to "calculate backwards": Since 3 birds flew from the first tree there had to be 6 + 3 = 9 birds in the beginning. An additional 7 birds came to the third tree, which means that initially there had to be 24 – 7 = 17 birds. Therefore, 16 birds must have been sitting originally on the second tree (42 – 9 – 17 = 16). One can do the following to check: 3 birds came to the second tree and 7 left it—which means that after the flight there had to be 4 birds less than before, and 12 + 4 = 16.

The second method does not contain any mathematics unknown to 12-year-old students, so it is, at least arithmetically, a little bit more economical. But in this case, one has to recognize first that the simple situation of the birds after the flight makes a better basis for the solution of the problem.

Example 2.5

Two athletes—Anton and Ben—start running from opposite ends (*A* and *B*) of a long, straight alley (Anton starts from *A*, Ben starts from *B*). They each run at a constant speed and meet 800m away from the nearest starting point. They then continue running, and, after reaching the other end of the alley, they run back immediately and meet each other 400 m away from the other starting point. How long is the alley (see Figure 10.2)?

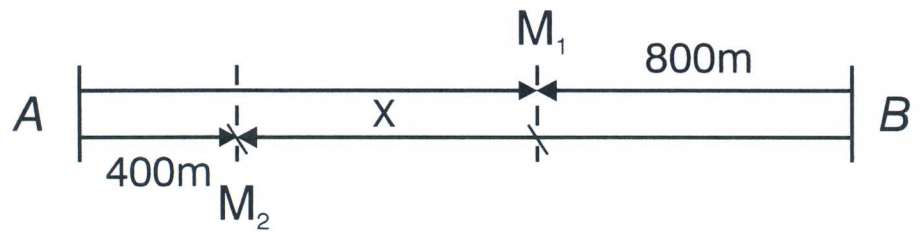

Figure 10.2

Again, we want to present two solutions that vary from each other in a similar way as in the above mentioned example. Maybe it should be emphasized that a sketch of the situation is very important for both solutions. A great part of the solution is already here in a suitable depiction.

Solution 1. The described situation may pose quite a dilemma for most students, as only very few things seem to be given and it may be not easy to find a "statement."

Let us suppose that Ben, who is starting from B, is the slower one. Then the first meeting point M_1 must be nearer to B than to A; $M_1B = 800$ m and $M_2A = 400$ m (Figure 10.2). It seems reasonable to let x be the length of the unknown piece, $\overline{M_1 M_2}$ of $\overline{AB}$, the more so since the length of $\overline{M_2 M_1}$ is required. However, neither of the two runners' speeds is known, nor are any times! As both runners started at the same time, it takes them the same time to get to the two meeting points, their
3 running times from the start to M_1 must be equal, and likewise from M_1 to M_2! Let v_1 be Anton's average speed and v_2 Ben's; then we get the following equations:

$$\frac{400+x}{v_1} = \frac{800}{v_2}$$

$$\frac{800+800+x}{v_1} = \frac{x+400+400}{v_2}$$

But there are three variables in these equations! It is not possible to get three values from two equations, is it? Did we not think of all the information? Is this problem really solvable?

Indeed, in general, it is not possible to extract three variables (here x, v_1, and v_2) from two equations, and we did not forget anything. It is immediately clear that the runners' speeds cannot really be defined by the information given above: If both runners, for instance, were running twice as fast, then they both would be at the meeting points in half the time, but the meeting points would not change. So one can expect only a statement about the relationship of both speeds (but this is not really asked here.)

We surely cannot compute all three variables, but perhaps we could find at least one—how about the value of x? This is indeed possible, by dividing the first equation by the second equation (no occurring terms are zero); both speeds then disappear from the equation and we get an equation in x:

13
$$\frac{400+x}{1{,}600+x} = \frac{800}{800+x}.$$

So we get the quadratic equation $320{,}000 + 1{,}200x + x^2 = 1{,}280{,}000 + 800x$, with the only positive solution being $x = 800$m. So the alley is 400m + 800m + 800m = 2,000m long. The relationship of the two speeds we get by inserting:

$$\frac{1{,}200}{v_1} = \frac{800}{v_2} \Rightarrow v_1 = \frac{3}{2}v_2$$

or, alternately, $v_1 : v_2 = 3 : 2$.

Solution 2. This solution does not need quadratic equations or, indeed, any equations at all, so it could be seen as particularly "elegant" in that way. Arithmetically, it contains, again, almost only elementary mathematics, but one has to see the situation in a suitable way:

"At their first meeting both sprinters together ran the whole length of the way exactly once." Have we already gained something by this? Not completely, but we've obtained a basis for the next steps! What is the situation at the meeting point M_2? At M_2, they have been running exactly three times the length of the alley (see Figure 10.2)! Now, as both sprinters run at a constant speed, not only has the total sprinted distance been tripled, but so have the individual distances run by each of the sprinters. Therefore, Anton and Ben are running from the start to M_2 three times the distance just to M_1. That means that Ben has to run a distance, from the start to M_2, of 3×800 m = 2400 m. Of that distance, he was running already 400 m back toward his starting point. Therefore, the length of the alley has to be 2400 m – 400 m= 2000 m! 3

2.3. *Problems That Do Not Rely on Equations—Problems of Order*

Although "text problems" and translating colloquial texts into the language of mathematics refer, respectively, in most cases to setting up and solving equations, not all described situations have to lead to equations. The latter kind of problems are extremely suitable for classes, as they stimulate the activity of the students and show the great variety of ways of thinking mathematically. Two connected examples follow.

Example 2.6

GEN

There are six patients (*A*, *B*, *C*, *D*, *E*, and *F*) sitting in the waiting room of a dentist. The dentist knows (on the basis of previous treatments) what needs to be done for these particular patients and can therefore guess how long it will take him to treat each one. The treatment of *A* will last about 15 minutes; for *B* he estimates 30 minutes; for *C* and *D* approximately 10 minutes each; for *E* about 20 minutes; and for *F* only 5 minutes. The dentist wants to keep the total waiting time of all patients as low as possible and is thinking of an order in which to see the patients. In this sense, is there an optimal order?

If he calls them, for instance, in the order *A*, *B*, *C*, *D*, *E*, *F*, so *A* does not have to wait at all, *B* would have to wait 15 minutes, *C* 45, *D* 55, *E* 65, and *F* 85. All together, this makes 265 min = 4 h 25 min. Are there more favorable orders? Is there, perhaps, even a most favorable one? If so, what is it? Is there a general rule for such problems? 14

Students could (should) try some orders, but it will hardly be possible to register all 6! = 720 possibilities in a systematic way.

How do the single waiting times and the total waiting time come about? Already in the example above it can easily be seen: If a patient Y has his turn after patient X, then the waiting time of Y obviously consists of the sum of the waiting time of X and the treatment time of X! The patients $P_1, P_2, \ldots, P_n$ have the times for their treatments $t_1, t_2, \ldots, t_n$. If they take their turns in that order, the following waiting times emerge:

13

$$
\begin{aligned}
&P_1\colon 0\\
&P_2\colon t_1\\
&P_3\colon t_1 + t_2\\
&\ldots\\
&P_{n-1}\colon t_1 + t_2 + \ldots + t_{n-2}\\
&P_n\colon t_1 + t_2 + \ldots + t_{n-2} + t_{n-1}.
\end{aligned}
$$

We add to get the total waiting time T:

$$T = (n-1)t_1 + (n-2)t_2 + \ldots + 2t_{n-2} + t_{n-1}.$$

The time for the treatment of the patient who is being treated first (t_1) is multiplied by the biggest factor $(n-1)$, thus giving it the "greatest weight" for the total waiting time, as this particular time has to be waited by all following patients. Now it can immediately be seen that if T shall become as small as possible, then the biggest factor $(n-1)$ has to be multiplied by the smallest possible value of t, $(n-2)$ has to be multiplied by the second smallest, and so on.

Result. The total waiting T time is the least if each time the patient with the shortest treatment time—among all still-waiting patients—has his turn first. The (one) optimal order for the initially asked problem is therefore $F \to C \to D \to A \to E \to B$, with a total waiting time of $T = 5 \times 5 + 4 \times 10 + 3 \times 10 + 2 \times 15 + 20 = 145$ min = 2 h 25 min.

So there is a waiting time saving of 2 hours compared with the alphabetical order—a considerable difference!

Remark. If there are patients with the same time for treatment, there is, of course, more than one optimal order. In our problem, the order $F \to D \to C \to A \to E \to B$ would also be an optimal one, as C and D have the same treatment time. We see that there may be more than one optimal solution to a problem, so, often, one cannot speak of *the*, but of *one*, solution.

Remark. Especially for young students, it is very important to learn that problems can have more than one solution. On the other hand, it is difficult to create such problems in an interesting way before equations of second or higher degree are taught. But certainly, this is too late!

An analogous problem, where one probably would have acted correctly intuitively, is the following:

As a winner of a lottery one can take bills from three piles of bills in the denominations \$100, \$50, and \$10. From any pile you choose first 10 bills, then from another pile 5 bills, and finally 1 bill from the third pile. To get as much money as possible, everybody would probably (without much thought) take the 10 bills from the \$100 pile, the 5 bills from the \$50 pile, and 1 bill from the \$10 pile. The sum of the winnings is maximal if the biggest factor (10) is multiplied by the largest value (\$100). The sum of the winnings corresponds here with the total waiting time (but as big as possible instead of as small as possible), and the values \$100, \$50, and \$10 correspond to the waiting times. The order of the piles (with the fixed order of the taking of the bills $10 \rightarrow 5 \rightarrow 1$) corresponds to the order of the patients.

Now a second problem of that kind (generalization!).

Example 2.7

A senator has to welcome five special-interest groups for brief talks (separately, one after the other). The members of the groups are already in the waiting room, and every group already has its fixed time for the consultation (compare with Table 10.1). In what order should the senator call in the groups in order to keep the total waiting time of all the individual members as low as possible?

GEN

The situation is again similar to the one at the dentist's office (but now with groups rather than individuals).

Table 10.1

Group	*Number of Members*	*Consulting Time*
1	4	20 min.
2	8	10 min.
3	5	30 min.
4	10	15 min.
5	6	25 min.

6

14

Solution 1. If group 1 (consisting of 4 members) is, for instance, in consultation with the senator for 20 minutes, it can be considered the same as if each of the these 4 persons were there for 5 minutes ($20/4 = 5$). So it is possible to split the groups

in single persons and the group times into single times (times per person). The above example gives a solution to the problem of an optimal order of single persons: A person with the shortest possible consultation time has to take his or her turn first. So the times per person determines the single order and in the sum, therefore, also determines the order of the groups, because in an optimal order of the individuals, the members of a group would certainly have their turn for the consultation one after another according to the above principal and could be seen in this respect as a group again. In other words, to call all persons with the currently shortest consulting time into the room is the same as calling in the whole group in its entirety first!

The consulting time per person in the single groups is indicated in Table 10.2:

Table 10.2

Group	1	2	3	4	5
Time per Person (min.)	5	1¼	6	1½	4⅙

Therefore, the optimal order for the groups, according to the increasing times per person, is $G_2 \rightarrow G_4 \rightarrow G_5 \rightarrow G_1 \rightarrow G_3$.

3 *Solution 2.* The following would be a slightly different way of dealing with the problem (for instance, in the case where the example above was not used). Let us reduce the problem first to the two groups G_1 and G_2. Is it better to take G_1 first or G_2?

17 If G_1 has its turn first, then the 8 persons of G_2 have to wait 20 minutes each (i.e., the consulting time of G_1), with a total waiting time of 160 minutes. If G_2 were to go in first, the 4 persons of G_1 have to wait 10 minutes each, with a total waiting time of 40 minutes, which is considerably less.

13 Let us now consider the general problem again: n different groups $G_1, \ldots, G_n$ with $g_1, \ldots, g_n$ members and the consulting times $t_1, \ldots, t_n$ are sitting in the waiting room. Let us assume they take their turn in the order $G_1, \ldots, G_k, G_{k+1}, \ldots, G_n$ to see the senator, and this results in a total waiting time T_1. What happens if any two consecutive groups are interchanged? If one interchanges, for instance, the two groups G_k and G_{k+1}, we get the order $G_1, \ldots, G_{k+1}, G_k, \ldots, G_n$ and a new total waiting time T_2. There is obviously no change in the waiting times of any of the groups not involved in this particular change ($G_1, \ldots, G_{k-1}$ and $G_{k+2}, \ldots, G_n$); nor does the combined waiting time that G_k and G_{k+1} wait together for $G_1, \ldots, G_{k-1}$ change. Let the sum of these (in every case) constant waiting times be S. Then T_1 consists of S and the waiting time of G_{k+1} for G_k: $T_1 = S + t_k g_{k+1}$

T_2 is derived analogously if G_{k+1} has its turn before G_k: $T_2 = S + t_{k+1} g_k$.

Under what conditions is T_1 smaller than T_2? Obviously, when $t_k g_{k+1} < t_{k+1} g_k$, which is equivalent to $\frac{t_k}{g_k} < \frac{t_{k+1}}{g_{k+1}}$.

So if t_k/g_k (the consulting time per member of the group G_k) is smaller than t_{k+1}/g_{k+1}, it is more favorable to ask the group G_k to the consultation first.

Also from this we can deduce in general: If it should occur in a certain order that the inequality $\frac{t_k}{g_k} > \frac{t_{k+1}}{g_{k+1}}$ holds true for any k, it is possible to reduce the total waiting time by exchanging the two groups.

Result. The order is optimal if and only if $\frac{t_1}{g_1} \leq \frac{t_2}{g_2} \leq \ldots \leq \frac{t_n}{g_n}$. So the result here as well is that the consulting time per person is decisive in finding the best order.

Especially for such problems, it is extraordinarily important to give the students enough time for presumptions and trials and to talk about all suggestions and arguments. When necessary, the students may be given little hints—the teacher should not disclose too much and should encourage students to work independently. The teacher, of course, has to lead or channel all problem-solving processes to a certain degree, since the students will probably often choose wrong approaches or will just not know what to do next. But this leading should not take the form of an especially elegant solution that appears from nowhere.

3. Problems in the Context of Special Mathematical Theories

This section will deal with some problems that illustrate the essence and the importance of certain mathematical themes and that contain easily recognizable "typical features" of these theories. Such (or similar) examples may sometimes even serve as an entry into certain mathematical topics if they are suitable for discovering the essentials of the new topic and if they motivate students to deal thoroughly with the particular issue. Such entry problems must not be too sophisticated and should show clearly that an increase in knowledge (e.g., precision, theoretical investigation, systematization, etc.) is necessary or at least possible.

3.1. A Problem Representing Different Theories

Example 3.1

"The Tower of Hanoi" (e.g., Haussmann, 1986; Stowasser & Mohry, 1977): A tower of n (e.g., $n = 3$) discs is built on one of three rods, with the discs getting smaller from bottom to top (see Figure 10.3). GAME

The task is now to move this "n-tower" from one rod to another while obeying two rules:

a. Only one disc is allowed to be moved from one rod to another each time.

b. A larger disc may never rest on a smaller one.

What is the minimal number $H(n)$ of single steps (moving of discs) for a tower of n discs?

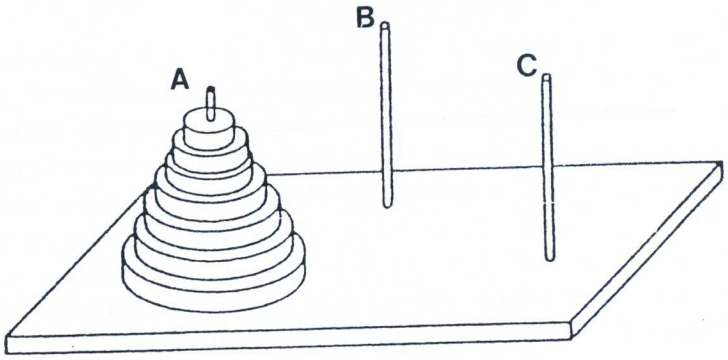

Figure 10.3 The tower of Hanoi

It is certainly a great advantage to have a few real trials with "model" towers of objects that get smaller (e.g., books, pieces of paper). By trial and error one can immediately make the following table for small n:

Table 10.3

2

n	1	2	3	4
$H(n)$	1	3	7	(15)

Many students will already begin having problems with $n = 4$, but for students with even a little interest, $H(4) = 15$ will not be too hard to work out. But now the question seems to be quite obvious, whether there is a "system," and whether it is possible to get $H(4), H(5), H(6), \ldots$ by a thought experiment (e.g., from the known values of $H(1)$, $H(2)$, and $H(3)$, respectively). Let us go back one more step: Is it already possible to get $H(3) = 7$ by consideration, without really experimenting? Put a different way: How can I move a 3-tower if I know how to move a 2-tower?

3

18

One can generally determine an answer easily from the concrete experience in experimenting: To move a 3-tower from rod A to rod C, one first has to move the upper 2-tower to rod B, then the lowest (largest) disc from A to C, and finally the 2-tower "stored" in B to C (and that by employing the same scheme used to move it before from A to B; see Figure 10.3). As it is known how a 2-tower has to be moved and how many steps it takes—$H(2) = 3$—it follows that $H(3) = H(2) + 1 + H(2)$; $7 = 3 + 1 + 3$, as the experiment confirmed already.

Now the structure of the problem becomes clearly visible: The same principle must now be applicable for $H(4)$: $H(4) = H(3) + 1 + H(3) = 2 \times H(3) + 1$. Analogously, to move a 4-tower, one has to move a 3-tower first, then the lowest disc, and finally again the 3-tower. So now it is not especially hard to discover: $H(n) = 2H(n-1) + 1$.

The value $H(n)$ is easy to determine if $H(n-1)$ is known, this in turn if $H(n-2)$ is known and so on, until one gets to a known $H(k)$. By this, one can determine step by step: $H(k+1) = 2H(k) + 1$, $H(k+2) = 2H(k+1) + 1, \ldots, H(n) = 2H(n-1) + 1$. 9

In this manner, the above table can be conveniently continued: The last number with which it seems sensible to conduct the experiment in reality is $n = 6$, because starting at $n = 7$ the number of necessary single steps is already very high.

Table 10.4

n	1	2	3	4	5	6	7	8	9	10	...
H (n)	1	3	7	15	31	63	127	255	511	1,023	...

One can see that this example very naturally leads to the principle of recursion or iteration. So it can be a gateway to the accompanying theory if the student has not heard of it. Another theory demonstrated here would be that of the difference equations (such as linear difference equations with constant coefficients), which are increasingly important for the description of dynamic systems, especially those of first order ($a_n = ka_{n-1} + d$) and perhaps those of second order ($a_n = k_1a_{n-1} + k_2a_{n-2} + d$). ALG

The principle of proof by mathematical induction can also be illustrated and even motivated by this problem. Generally, it seems that proofs by induction especially are often introduced without any natural motivation, without "discovery" and through a relatively dry lecture that doesn't help students really understand the necessity or the principle itself! ALG 18

For big n (e.g., $n = 50$), of course, this step-by-step calculation of the values of $H(n)$ is relatively clumsy (at least without a programmable calculator or computer). So the question arises whether it is possible to express $H(n)$ directly by a formula without having to start at $H(2)$ or $H(3)$ and coming forward only gradually. For this we should look again at Table 10.4. The powers of 2 (2, 4, 8, 16, 32, 64, 128, . . .) are very significant and impressive numbers, so one will soon notice that the values of $H(n)$ are always 1 smaller than 2^n. Therefore, the assumption : $H(n) = 2^n - 1$ seems justified. According to Table 10.4, it is true at least for $n \leq 10$.

Now, what does the case $n = 11$ look like? We know $H(10) = 2^{10} - 1$, and we further know that $H(11) = 2(2^{10} - 1) + 1 = 2 \times 2^{10} - 2 + 1 = 2^{11} - 1$.

So our assumption is true as well for $n = 11$. We proved the validity of the formula for $n = 11$ by the validity for $n = 10$. For $n = 12$ we therefore get $H(12) = 2H(11) + 1 = 2(2^{11} - 1) + 1 = 2 \times 2^{11} - 2 + 1 = 2^{12} - 1$.

One can already see that the validity of the formula is "inherited" from one number to another (one could also use further concrete examples). It is already perceptible, from these concrete examples, that this hereditary character does not depend on the particular choice of n (e.g., $n = 11$ or $n = 12$). We could repeat the above step to prove thereby the validity of the formula for any fixed n. But it is much more convenient to show this hereditary character in general (i.e., detached from concrete numbers): We want to show the heredity of the formula for all $n \in N$. Let us assume the validity of the formula $H(n) = 2^n - 1$ for any n. It then follows that $H(n + 1) = 2H(n) + 1 = 2(2^n - 1) + 1 = 2^{n+1} - 1$, and this is the validity of the formula for the next number $n + 1$.

But if a formula is true for $n = 1$, and if the validity for any natural number ensues from the validity for the previous natural number, then the formula must obviously be true for all natural numbers:

$$H(1) \rightarrow H(2) \rightarrow H(3) \rightarrow H(4) \rightarrow H(5) \rightarrow \ldots$$

We believe with this example, students could discover the theories of recursion and complete induction almost by themselves in a genetic way, and could therefore perhaps understand them better (see Stowasser & Mohry, 1977, 7ff).

Other possibilities of proving the validity of $H(n) = 2^n - 1$ would be, on the
3 one hand, through the sum formulas of geometric series (another theory) or, on the other hand, by the following consideration:

We know the recursion equation $H(n + 1) = 2H(n) + 1$ with $H(1) = 1$ is true. From that we get, step by step:

$$\begin{aligned}
&H(2) = 2 \times 1 + 1 = 2 + 1 = 2 + H(1)\\
&H(3) = 2(2 + 1) + 1 = 2^2 + 2 + 1 = 2^2 + H(2)\\
&H(4) = 2(2^2 + 2 + 1) + 1 = 2^3 + 2^2 + 2 + 1 = 2^3 + H(3)\\
&\ldots\\
&H(n + 1) = \ldots = 2^n + 2^{n-1} + \ldots + 2 + 1 = 2^n + H(n).
\end{aligned}$$

By equating both terms for $H(n + 1)$, one gets $2H(n) + 1 = 2^n + H(n)$, from which immediately follows $H(n) = 2^n - 1$.

3.2. *Simple Tasks to Be Solved by Graphic Representation*

A special device for the solution of some problems is graphic representations. Drawing sketches often initiates and supports the process of problem solving and takes it in the right direction (see examples 2.2 and 2.5). An example of a graphic illustration making an important contribution to the solution comes in network technique.

Network technique is a relatively young method (developed in the late 1950s) for better planning, controling, coordinating, and supervising of bigger enterprises

or projects. One of the most essential aims of this technique is to determine the earliest possible completion date of a project, which consists of several "partial tasks."

Mathematics can be seen in a certain way as a language, and its special structure can contribute decisively to the solution of many problems. Its means are, among others, variables, equations, and graphics (illustrations). Graphic presentations of all kinds especially serve to describe situations as they can also be described by variables, equations, formulas, and the like; but graphic presentations can often be realized better, more easily, and more quickly than texts or representations with variables. "A picture *is* worth a thousand words." 6

The network technique is an excellent opportunity to make these graphic aspects clear. Here, not only are graphic representations worked out, interpreted, and described, but there is an essential additional factor: One can also see this technique as an "intelligent" graphic representation in the sense that logical connections, formalized relations, quantitative connections, and the like are illustrated in such a way that the representation already gives a huge part of the solution. Here is a very simple example that can be especially suitable as an entry-level problem.

Example 3.2

A small construction project is divided into certain stages $P_1, P_2, \ldots, P_6$ that are achieved by the completion of single jobs. P_1 will be the beginning stage and P_6 the end stage. The notation $\overline{P_2 P_5} = 9$ means that P_2 is a precondition for P_5, and the work that has to be done "between" P_2 and P_5 takes at least 9 days (time units). The complete information on the whole project is given by: ALG

$$\overline{P_1P_2} = 4;\ \overline{P_2P_3} = 6;\ \overline{P_2P_5} = 9;\ \overline{P_2P_4} = 8;\ \overline{P_3P_5} = 7;\ \overline{P_5P_6} = 3;\ \overline{P_4P_6} = 6.$$

The lengths of the single working processes $P_i \rightarrow P_j$ is based on experience. What is the minimal length of time for the entire project?

Solution. The given information about the single $\overline{P_iP_j}$ contains information both about which stages have to be reached before others and about how long each single working process lasts. The graphic translation of the above information leads to a netlike representation (Figure 10.4) called a "network," in which the single working process is marked as an arrow between the single stages. Such an representation gives a good survey of the situation and allows, in simple cases, a solution by trial and error. (See Chapter 15.) 11

Since, for completion, all processes have to be done, one has to look for the longest path from P_1 to P_6 to get the shortest total time. The other processes can then easily be "fit in." This may present a slight paradox for some students. They intuitively might have looked for the shortest path, but it is not difficult to see that the jobs that last longer could not be completed then. By trying all possible paths from P_1 (start) to P_6 (end)—adding all the working times along the paths—in simple "networks" such as the one in Figure 10.4, one will quickly see that here the "critical path" is $P_1 \to P_2 \to P_3 \to P_5 \to P_6$ ("critical" only because the jobs along this path obviously must not be delayed to prevent an extension of the shortest total time). The processes $P_2 \to P_5$ and $P_2 \to P_4 \to P_6$ can be fit in parallel fashion, as they are shorter. The shortest time for the realization of the project (e.g., construction time) is therefore 4 + 6 + 7 + 3 = 20 time units.

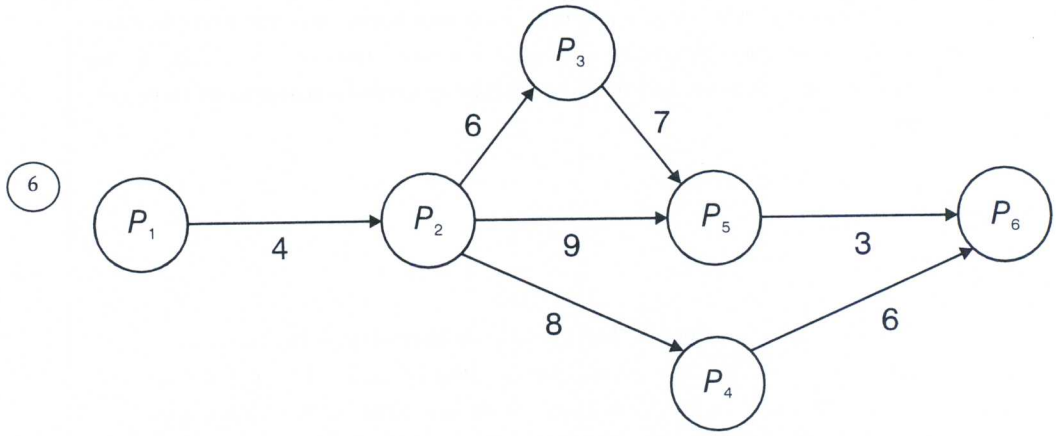

Figure 10.4 Network of a small project

Remark. If students have never been confronted with this kind ot task, even simple ones such as the above can pose a problem in the sense that students have to look for a (for them) new kind of solution.

Also, in slightly more complicated "networks" (e.g., Figure 10.5), it is possible to find one solution by trying different paths. As this example shows, there may be several solutions (i.e., the critical or longest paths). In the network of Figure 10.5, the two paths $P_0 \to P_1 \to P_5 \to P_7$ and $P_0 \to P_3 \to P_4 \to P_7$ are both critical ways with a total length of 15 time units each (see Reichel et al., 1989-1992, p. 171).

For the solution of very complex building projects, of course, this method of trial and error will not be sufficient anymore. Some algorithms have already been developed to determine the critical (longest) paths with the help of highly efficient computers. These algorithms are not at all so complicated that they could not be dealt with in school. There is really a theory behind such tasks, and even students can gain a first insight into it.

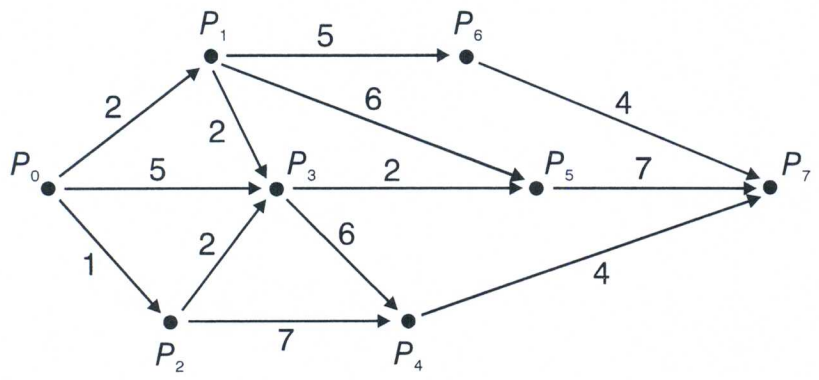

Figure 10.5 Example of a network

3.3. *Problems Concerning Rational and Irrational Numbers*

We do not know if all students develop the right idea of the importance and the nature of irrationality by the current examples in school (e.g., proof of the irrationality of $\sqrt{2}$). What can it mean or what effect does it have if a number cannot be expressed as a fraction (i.e., the relation of two whole numbers)? The following problems, along with the solutions, will show some surprising results.

Example 3.3

In a system of coordinates, a point with integer coordinates is called a *lattice point* (the point [4, 3], for instance, would be a lattice point, and the point [0.5, 2] would not).

Is there a straight line through the origin of a two-dimensional system of coordinates that does not contain any lattice point (Figure 10.6)? For reasons of symmetry, one can confine attention here to positive coordinates.

ALG

6

Of course there are many straight lines that do not contain a lattice point in the shown sector of Figure 10.6, but whether they do not meet a lattice point "anywhere" is surely an exciting question!

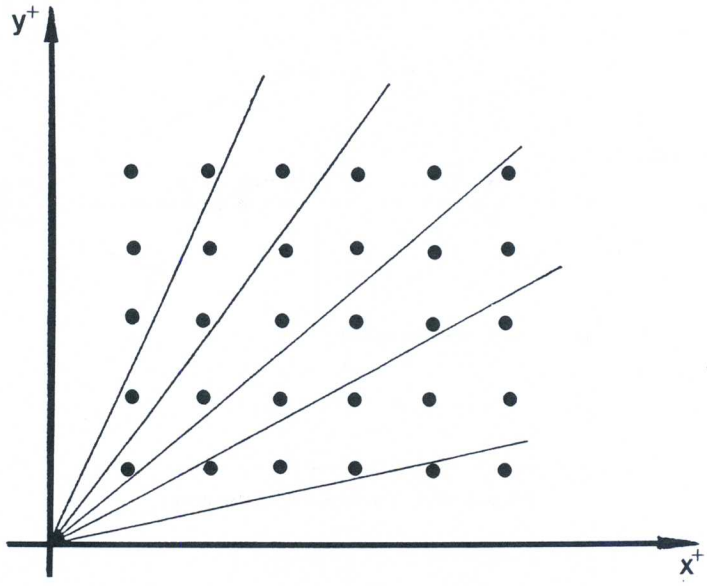

Figure 10.6 Lines and lattice points

Another formulation could be: Is it possible to fire from the origin an idealized "arrow" (with diameter zero) through a "forest" with "trees" at the lattice points (also with diameter zero) without ever hitting a "tree" (the "arrow" may eternally maintain its original direction)?

Solution. If a straight line contains a lattice point with the coordinates (m, n) $m, n \in N$, it has the slope n/m. This means that if it is not possible to express the slope of a straight line as a quotient of two natural numbers, then it cannot contain any lattice points. The inverse is, of course, true as well: If a straight line has the slope n/m, it will contain a lattice point, such as (m, n). So only the straight lines with
13 irrational slopes do not contain lattice points, and for this there are many more possibilities than for straight lines with rational slopes. Although there is an infinitude of straight lines with both rational and with irrational slopes, one can formulate: "The majority of straight lines through the origin do not contain lattice points" (noncountable infinitude in contrast to countable infinitude of straight lines with lattice points).

To continue, it could be shown: If a straight line through the origin contains one additional lattice point (other than [0, 0]) it must contain an infinite number of them! What is the distance between two neighboring lattice points?

Solution. If (a, b) is a lattice point, then $(2a, 2b)$, $(3a, 3b)$, . . . are, of course, lattice points, too. These are neighboring points if a and b are relatively prime. In this case we get, for the distance, $\sqrt{a^2+b^2}$.

Example 3.4

Take any circle and mark a point P_0. Then translate (rotate) P_0 by the angle α to obtain the point P_1. Repeat the process to obtain P_2, P_3, P_4, . . . (see Figure 10.7). One of the points P_n may be equal to point P_0 again. In that case, our sequence of points is finite. On the other hand, the sequence obtained by the process described can be infinite, too. Then it consists of pairwise different points, of course. Now the question arises: For which angles α do we get finite sequences, and for which α infinite ones? In other words, for which α is there an index n such that $P_n = P_0$? ALG

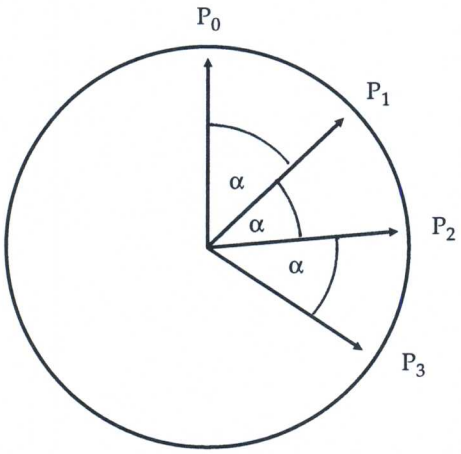

Figure 10.7 Finite or infinite sequence

Solution. If $P_n = P_0$, for some n, some multiple of α is a multiple of 360°, too; in other words, there are natural numbers k and l so that the following is valid:

$$360 \times k = \alpha \times l \iff \alpha = \frac{360k}{l}. \quad (13)$$

This means α has to be rational. Conversely, if α (measured in degrees) is rational, the sequence $P_0, P_1, P_2, P_3, \ldots$ is finite. Since, if $\alpha = a/b$ $(a, b \in N)$, we have $(360° \times b) \times \alpha = a \times 360°$. In other words, $P_0 = P_{12}$ for $n = 360° \times b$. In conclusion, $P_0, P_1, P_2, P_3, \ldots$ is infinite if and only if α (measured in degrees) is irrational.

Example 3.5

GEOM

A circle K_1 with the radius r_1 touches another circle K_2 with the radius r_2, and the points that touch are called $P_0 \in K_1$ and $P \in K_2$, respectively (see Figure 10.8a). By rolling along the circumference of K_1, the point P describes an arched curve called an *epicycloid* (see Figure 10.8b), where the "next" points that touch (when P meets again exactly with K_1) are called $P_1, P_2, \ldots$ Under which conditions of r_1 and r_2 will the curve of the point ever close? (When will the point P ever return to the starting point P_0? When is there some $r \in N$ with $P_n = P_0$?)

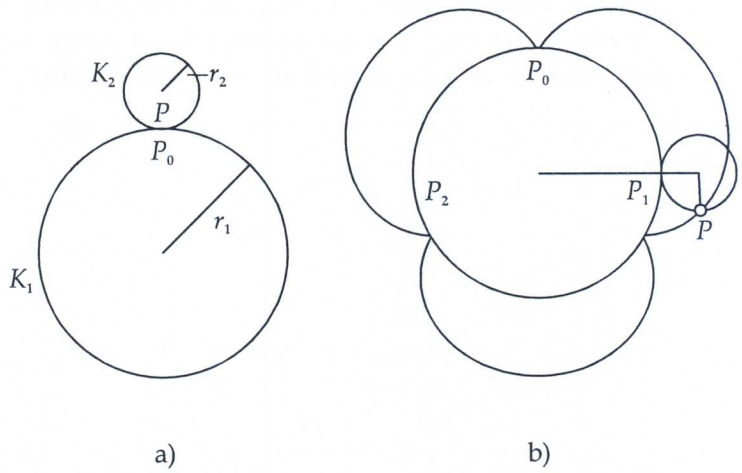

Figure 10.8 Rolling of a circle—"epicycloids"

Solution. Here one could rush to the conclusion: The curve will be closed only if r_1 and r_2 have rational values, otherwise not. The curve will obviously be closed if and only if both circumferences (u_1 and u_2) have a common multiple—that is, if there are natural numbers k and l with $(2r_1\pi) \times k = (2r_2\pi) \times l$. This is equivalent to
13 $r_1 \times k = r_2 \times l$ or, alternately, $r_1/r_2 = l/k$. Therefore, it is not the irrationality of r_1 or r_2 individually that is decisive here, but that of u_1/u_2 or r_1/r_2. If, for instance, $r_1 = 3\pi$ and $r_2 = \pi$, then the curve obviously closes exactly after one "round" of K_1 or, equivalently, after three resulting "arches" (e.g., see Figure 10.8b), though both radii have irrational values.

Result. If the fraction r_1/r_2 is irrational, then the curve of P will never close. If r_1/r_2 is rational, so that $r_1/r_2 = m/n$ (m and n relatively prime), then the curve will close after exactly n rounds of K_1, or equivalently, after m arches.

To better appreciate the full flavors of rationality, we shall consider a few more problems.

Example 3.6

Are there positive numbers $a, b \in Q$, but $a, b \notin N$, so that $a^b = c \in N$? ALG

Most students will know the following proposition: The square root of a natural number is either again a natural number or an irrational number. The same holds true for the kth root instead of the square root (to be proved analogously).

Solution. There are no such numbers a and b. 13

Indirect proof: Let $a = m/n$ ($a \notin N$) and $b = k/l$ ($b \notin N$) then the above condition states

$$\left(\frac{m}{n}\right)^{\left(\frac{k}{l}\right)} = c \ (c \in N).$$

This yields:

$$\left(\frac{m}{n}\right)^k = c^l \text{ or, equivalently, } \frac{m}{n} = \sqrt[k]{c^l}$$

As $c^l \in N$, then, according to the above proposition, $\sqrt[k]{c^l}$ is either a natural number itself (contradiction to $m/n = a \notin N$) or an irrational number (contradiction to $m/n = \sqrt[k]{c^l}$). That is why such numbers cannot exist.

Example 3.7

Are there $a, b \in R$ but $a, b \notin Q$, so that $a^b = c \in Q$ (cf. Heinze, 1993, 147f.)? ALG

Solution. $\sqrt{2}$ is surely an irrational number. Consider, for instance, $\sqrt{2}^{\sqrt{2}}$—one of the simplest expressions of the demanded form—and let us ask if this value is rational or not. If this value is rational, then we have already found one example 13

and have to answer the question posed with Yes; if not, we have to continue working. So let us assume that $\sqrt{2}^{\sqrt{2}}$ is not rational and continue trying. We form, for instance,

$$(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^{2} = 2$$

and see that the result, 2, is not only rational but also a natural number, though the exponent ($\sqrt{2}$) is irrational and (according to the assumption) the base ($\sqrt{2}^{\sqrt{2}}$) is irrational, too.

Result. If $\sqrt{2}^{\sqrt{2}}$ is rational, then we have to say "yes" to the question posed; if it is irrational, then with $(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}$ we have found an example of the demanded form and have to answer this question as well with Yes. Even without knowing whether $\sqrt{2}^{\sqrt{2}}$ is rational or irrational, we have found an example in every case.

Analogous examples would be $(\sqrt{5}^{\sqrt{2}})^{\sqrt{2}}$ or $(\sqrt[3]{2}^{\sqrt{3}})^{\sqrt{3}}$.

3.4. Problems Concerning Complex Numbers

The following problems have the theory of complex numbers as a background and contain some surprising results. Especially surprising moments in class often promote motivation as well as fascination, excitement, and interest. But it should be emphasized that we are dealing here with problems that will find successful application in more advanced classes. The first problem is a continuation of example 3.7.

Example 3.8

ALG

Are there $a, b \in C$, but $a, b \notin R$, with $a^b = c \in R$?

13

Students generally learn Euler's formula $e^{i\phi} = \cos\phi + i\sin\phi$ without an exact foundation by expansion into series. But the exact derivation and the knowledge of all problems around the exact definition of the complex exponential function are not really necessary for the following considerations.

It is quite simple to insert some values into Euler's formula and just wait with excitement for the result. For $\phi = \pi$ one can already get the first interesting result: $e^{i\pi} = \cos\pi + i\sin\pi = -1$.

The numbers e and π are irrational (even transcendental), $i \times \pi$ is a complex number, and $e^{i\pi}$ is nevertheless an integer—really surprising! For $\phi = 2\pi$ one gets $e^{2\pi i} = 1$—a natural number.

Now let us take $\phi = \pi/2$ and substitute it; we get $e^{i\frac{\pi}{2}} = i$—also an interesting identity. The simplest power where base and exponent are not real numbers is i^i. We get

$$i^i = \left(e^{i\frac{\pi}{2}}\right)^i = e^{i^2\frac{\pi}{2}} = e^{-\frac{\pi}{2}} = \frac{1}{\sqrt{e^{\pi}}},$$

a real number, which is expressed again by the numbers e and π, which are extremely important for mathematics in general.

With help of the Euler's formula we can also get an insight to the periodicity of the complex exponential function and can find as well solutions to problems like the following.

Example 3.9

Does the equation $1^x = 3$ have a solution in C? ALG

As seen already, Euler's formula yields, for $\phi = 2\pi$, the length of the identity $e^{2\pi i} = 1$. Since the functions sin and cos are periodical with a period 2π, the formula also holds true for $\phi = 4\pi, 6\pi, \ldots, 2k\pi, \ldots$: 13

$$e^{2k\pi i} = \cos 2k\pi + i \sin 2k\pi = 1 \ (k \in Z)$$

We see the equation $e^x = 1$ has not only the solution $x = 0$ but even an infinitude of solutions $x = 2k\pi i$ $(k \in Z)$. Now to the equation of the posed problem $1^x = 3$, which can be described—as we have just seen—in the following way, too:

$$(e^{2k\pi i})^x = e^{ln3} \Leftrightarrow e^{2k\pi i\, x} = e^{ln3},$$

from which immediately follows:

$$x = \frac{ln3}{2k\pi i} = -\frac{ln3}{2k\pi} \times i \ (k \in Z).$$

The given equation does not have a solution in R, but in C it even has infinite solutions, so this is an especially interesting proof of the "mathematical abundance" of complex numbers. Of course, one cannot carry on complex analysis in school, but a little insight into the possibilities of mathematical extension by complex numbers seems to us very interesting for students and beneficial to their motivation, especially in advanced courses.

Analytic geometry also offers some surprises and good opportunities to incorporate complex numbers into certain considerations.

Example 3.10

ALG

Given any circle and a point P outside of it, we easily can find the equation of the tangent through P. But what if P lies inside the circle. Is there still a tangent? Formally, the point P can be inserted into the tangent-equation. But is there any meaning in doing so?

13

It surely needs no explanation that this is not possible in the range of real numbers. But in the realm of complex numbers, even this is possible! We want to try, for instance, the point with coordinates (0, 0), the center-point of the circle $x^2 + y^2 = r^2$, to put formally/arithmetically the tangents on that circle. From analytical geometry we know that the line $y = kx + d$ will touch the circle $x^2 + y^2 = r^2$ if and only if $r^2(1 + k^2) = d^2$.

As the tangent shall be put on the circle from the point (0, 0), $d = 0$, and because of $r \neq 0$ it follows: $1 + k^2 = 0$, so $k = \pm i$.

Therefore, the equation of the tangents is $y = \pm ix$. As expected, these are imaginary lines, purely complex ones. These straight lines have some other surprising properties as well.

Which straight lines are orthogonal to $y = ix$ and $y = -ix$, respectively?

It is well known that two straight lines are orthogonal if for their slopes k_1 and k_2, the following is valid: $k_1 k_2 = -1$. For the straight line $y = ix$, $k_1 = i$; what is wanted now is a k_2 with $ik_2 = -1$ (i.e., $k_2 = i$). Therefore, the straight line that is orthogonal to $y = ix$ is $y = ix$ itself (!)—and this is not a misprint; in the range of complex numbers, there are straight lines that are orthogonal to themselves (analogously: $y = -ix$).

Take any pair of points P_1 and P_2 of the line $y = ix$ (or $y = -ix$, respectively). What distance will they have?

$P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ will be two points on the line $y = ix$. The length (Euclidean distance) of segment $\overline{P_1P_2}$ always equals zero:

$$(P_1P_2)^2 = (x_2 - x_1)^2 + (ix_2 - ix_1)^2 = (x_2 - x_1)^2 + i^2(x_2 - x_1)^2 = 0. \text{ (Since } i^2 = -1)$$

For a successful treatment of these topics, at least two preconditions have to be fulfilled:

1. The performance of the class must be at a relatively high standard, because the processes of abstraction that have to be executed are not really elementary.
2. The teacher has to be excited by the surprising results and also has to be able to transmit this excitement (tension) to the "positive attitude of expectation" of the students.

3.5. *Problems Concerning Sums of Natural Numbers*

The following problems refer to the field of "arithmetical series," especially to the formula

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

This formula in particular is often practiced only by using stereotypical and recipelike tasks. The following problems are a little more multifarious.

Example 3.11

Calculate the sum of all natural numbers up to 300 that are divisible by neither 8 nor 6. ALG

This is a typical task that can easily be divided into subtasks. First one could calculate the sum of all natural numbers from 1 to 300:

$$1 + 2 + 3 + \ldots + 299 + 300 = \frac{300 \times 301}{2} = 150 \times 301 = 45{,}150. \quad (13)$$

Now the sum of those numbers that are divisible by 8 or 6 is to be subtracted. The sum of all numbers (≤ 300) that are divisible by 8 is

$$8 + 16 + 24 + \ldots + 296 = 8(1 + 2 + 3 + \ldots + 37) = 8 \times \frac{37 \times 38}{2} = 5{,}624.$$

The sum of all numbers (≤ 300) that are divisible by 6 is:

$$6 + 12 + 18 + 24 + \ldots + 294 + 300 = 6 \times \frac{50 \times 51}{2} = 7{,}650.$$

If one now forms the sum 5,624 + 7,650 = 13,274, then one gets the sum of all numbers that are divisible either by 8 or by 6—or do we? No, there has to be a little correction: The number 24, for instance, was counted twice (see above), because it is divisible both by 8 and 6. We see that all numbers with this quality were counted twice, which means we have to subtract them once. Which numbers are these that are divisible by 6 and 8? It is not particularly difficult to see that they are the multiples of 24.

Therefore the sum

$$24 + 48 + 72 + \ldots + 240 + 264 + 288 = 24(1 + 2 + \ldots + 11 + 12) = 24 \times \frac{12 \times 13}{2} = 1{,}872$$

has to be subtracted from 13,274: 13,274 – 1,872 = 11,402. Now this number has to be subtracted from 45,150, and as the result we get 33,748.

Example 3.12

ALG

Beginning with 1, the natural numbers are added until one gets a three-digit number with three equal digits. How many numbers have to be added (cf. Baron & Windischbacher, 1990, pp. 14, 89f)?

3

Solution. One solution would be to simply add up step by step until one gets the demanded form. But we want to choose another method. We have to add n numbers, and this sum

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

13

may not exceed 999. By inserting, or by solving a quadratic equation, one can see easily that n may not exceed 44. We look, therefore, for an $n \leq 44$, so that

$$\frac{n(n+1)}{2} = xxx \qquad (1 \leq x \leq 9)$$

Now, unfortunately, the expression "xxx" is not very usable; therefore, we are looking for a more suitable and simpler expression of "xxx," such as $xxx = x \times 111$, and therefore

$$\frac{n(n+1)}{2} = x \times 111.$$

But the number 111 can be written even more basically as a product of primes $37 \times 3 = 111$; then we get

$$n(n+1) = 2 \times x \times 37 \times 3 \qquad n \leq 44; \qquad 1 \leq x \leq 9.$$

As 37 is a prime, either n or $n + 1$ must be divisible by 37. So with the condition $n \leq 44$, only the possibilities $n = 37$ or $n = 36$ remain. For $n = 37$ we get $37 \times 38 = 2 \times x \times 37 \times 3$, but this is not possible, as 37×38 is not divisible by 3. Therefore, the only remaining possibility is that $n = 36$. From $36 \times 37 = 2 \times x \times 37 \times 3$, it follows immediately that $x = 6$. The sum of the numbers from 1 to 36 is really 666.

The method used here (factorization of 111 as a product of primes) may look like a "trick." But when teaching problem solving, we always use such tricks that—and this is the most important item—turn out to be mathematical methods as soon as we recognize that there are common features and structures of problems corresponding to several kinds of "tricks." Factorization into a product of primes is a significant example of that!

Example 3.13

Determine the sum of the digit-sums of all natural numbers from 1 to 999. ALG

Solution. What does it mean, to determine sums of digit-sums of certain numbers? The digits are obviously counted (summed) as often as they occur in the single numbers; for example, the sum of the digit-sums of the numbers 123, 203, 125, and 52 is $(1 + 2 + 3) + (2 + 3) + (1 + 2 + 5) + (5 + 2) = 1 \times 2 + 2 \times 4 + 3 \times 2 + 5 \times 2 = 26$. 3 13

So we have to rephrase the question to ask how often the digits 1, 2, . . . , 9 occur in the numbers from 1 to 999! It is irrelevant how often the digit 0 occurs, as it does not contribute to the digit-sum. First of all it is clear (or could become clear after some concrete considerations) that the digits 1, 2, . . . , 9 occur equally often in the numbers from 1 to 999, as each number appears equally often in each place (hundreds, tens, and ones), but how often? Let us split the problem again. How often, for instance, does the digit 3 occur as a hundred-digit? Obviously 100 times (300-399). How often does the 3 appears as a ten-digit? In every "hundred section" 10 times (e.g., 130-139). Since there are 10 of these "hundred sections," the digit 3 occurs altogether $10 \times 10 = 100$ times as the ten-digit. As the one-digit, the digit 3 occurs as well 10 times in every "hundred section" (e.g., 103, 113, 123, . . . , 193), so altogether 100 times as well. We see that the digit 3 (as well as all the other digits, except 0, which occurs altogether 189 times) occurs 300 times altogether in the numbers from 1 to 999. The sum of the digit-sums is therefore

$$300(1 + 2 + \ldots + 9) = 300 \times \frac{9 \times 10}{2} = 13{,}500.$$

3.6. *An Interesting Problem Concerning the Divergence of the Harmonic Series*

The following example from Kranzer (1989, pp. 157, 223ff) should not be interpreted as a practice-oriented example. The wording here is only one aspect that should support the surprise from the result of this thought experiment and may contribute to motivation.

Example 3.14

ALG

A snail is sitting at the end A of a 1 km long and arbitrarily elastic rubber band. The snail has only one task in its infinitely long life: to creep at a speed of 1 cm/s toward the other end of the rubber band. After each second, the rubber band is stretched by 1 km by an invisible power (homogeneous stretching). Will the snail ever reach the other end B of the rubber band? If yes, how long will it take it?

13

Solution. Let us first consider how the length of the band changes in the course of the time. At each full second the band is homogeneously stretched by 1 kilometer. Let the time immediately before the completion of the nth second be n^-, and the one immediately after the completion of the nth second be n^+; between both the band has been stretched by 1 km. At the time n^-, the band is n km long, but at the time n^+, its length is $n + 1$ km, caused by the consistent homogeneous stretching of the whole band by the factor $(n + 1)/n$ "as quick as lightning." Also, the distance already covered by the snail is being increased.

Let us look first at the individual 1 cm pieces the snail covers during the 1st, $2^{nd}, \ldots, n^{th}$ second, in particular at the stretches that have an effect on these pieces.

The 1 cm piece of the first second is at the transition $1^- \to 1^+$ being stretched to its double length, at the transition $2^- \to 2^+$ to the $3/2$-fold length, . . ., in general at the transition $n^- \to n^+$ to the $(n + 1)/n$-fold. So at the time n^+ it has the length (in cm):

$$1 \times \frac{2}{1} \times \frac{3}{2} \times \frac{4}{3} \times \ldots \times \frac{n+1}{n} = \frac{n+1}{1}.$$

The 1 cm piece along which the snail creeps during the second second is in the transition $1^- \to 1^+$ not yet stretched as a covered distance; it is being stretched only at the transition $2^- \to 2^+$ to the $3/2$-fold length, at the transition $3^- \to 3^+$ to the $4/3$-fold length, . . . , in general at the transition $n^- \to n^+$ to the $(n + 1)/n$-fold length. It therefore has at the time n^+ the length (in cm):

$$1 \times \frac{3}{2} \times \frac{4}{3} \times \ldots \times \frac{n+1}{n} = \frac{n+1}{2}.$$

The 1 cm piece along which the snail creeps during the third second is being stretched only by the third stretching onward, for the length of this piece at the time n^+ we get analogously (in cm):

$$1 \times \frac{4}{3} \times \frac{5}{4} \times \ldots \times \frac{n+1}{n} = \frac{n+1}{3}.$$

Finally we look at the 1 cm piece that is covered in the nth second, which is stretched exclusively at the transition $n^- \to n^+$ with the stretching factor $(n + 1)/n$. So this piece has at the time n^+ a length of $(n + 1)/n$ cm.

The sum of all these stretched 1 cm pieces indicates the total distance of the snail from the starting point A. This distance is therefore at the time n^+:

$$\frac{n+1}{1} + \frac{n+1}{2} + \frac{n+1}{3} + \ldots + \frac{n+1}{n} = (n+1) \times \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right).$$

The band itself has, at the time n^+, a length of $n + 1$ km $= (n + 1) \times 10^5$ cm. The difference is the distance which the snail still has to cover:

$$(n+1) \times \left(10^5 - 1 - \frac{1}{2} - \frac{1}{3} - \ldots - \frac{1}{n}\right).$$

The negative fractions in the second parentheses are exactly the terms of the harmonic series, from which one knows that the series is divergent—that is, it grows beyond all limits. The numerical value of the series

$$\sum_{i=1}^{n} \frac{1}{i}$$

is for a very big n therefore also bigger than 10^5, and the value of the expression in the second parentheses is therefore negative; but if the distance that remains to be covered is no longer positive, then this means that the end has been reached already. It takes, of course, a very long time for this to happen. To get at least a very rough approximation of the order of magnitude of the necessary period of time, a formula for the approximation for a very large n could help. The exact foundation of the following limit-statement (Euler-Mascheroni's constant):

$$\lim n \to \infty \left[\ln n - \sum_{i=1}^{n} \frac{1}{i}\right] \approx 0.57721\ldots$$

will hardly be possible at school, but the teacher could just state it, without proof, to give an idea of the necessary period. For large n, we therefore have:

$$\ln n - \sum_{i=1}^{n} \frac{1}{i} \approx 0.57721\ldots \Leftrightarrow \sum_{i=1}^{n} \frac{1}{i} \approx \ln n + 0.57721$$

From the inequality $\ln n - 0.57721 > 10^5$, it follows immediately for n

$$n > e^{10^5 + 0.57721} \approx 10^{43,429.2} \approx \times 10^{43,429}.$$

This is the time in seconds the walking-tour takes the snail. But even expressed in years the number does not look much smaller—it would be approximately $\times 10^{43,421}$ years, an unimaginably huge number and an even more unimaginable period of time (according to today's knowledge, about 2×10^{10} years have passed since the Big Bang)!

References

Baron, G., & Windischbacher, E. (1990). *Österreichische Mathematikolympiaden 1970-1989: Aufgaben und Lösungen.* Innsbruck: Universitätsverlag Wagner.

Dürschlag, S. (1983). Problemlösen und Kreativität im Mathematikunterricht—Konzeption und Tendenzen in der polnischen Didaktikliteratur. *Der Mathematikunterricht, 29*(3), 46-70.

Engel, A. (Ed.). (1979). Internationale Mathematik-Olympiade. *Der Mathematikunterricht, 25*(1).

Haussman, K. (1986). Iteratives versus rekursives Denken beim Problemlösen im Mathematikunterricht. *Mathematica Didactica, 9*, 61-74.

Heinze, G. (1993). Förderung mathematisch-naturwissenschaftlich begabter Schüler in Sachsen am Beispiel der Mathematikausbildung am Gymnasium Dresden Blasewitz. *Didaktik der Mathematik, 21*(2), 140-153.

Kranzer, W. (1989). *So interessant ist Mathematik.* Köln: Aulis-Deubner.

Laub, J., Hruby, E., Reichel, H.-C., Litschauer, D., & Gross, H. (1985-1988). *Mathematik Arbeitsbuch* (4 vols.). Wien: Hölder-Pichler-Tempsky.

Polya, G. (1954). *Mathematics and plausible reasoning* (2 vols.). Princeton, NJ: Princeton University Press.

Polya, G. (1980). *Mathematical discovery* (2 vols.). New York: John Wiley.

Reichel, H.-C. (1991). Sprachschulung und Spracheinsatz im Mathematikunterricht. In H. Postel, A Kirsch, & W. Blum (Eds.), *Mathematik Lehren und Lernen* (pp. 156-170). Hannover: Schrödel.

Reichel, H.-C., Müller, R., Hanisch, G., & Laub, J. (1989-1992). *Lehrbuch der Mathematik* (Vols. 5-8). Wien: Hölder-Pichler-Tempsky.

Sell, R. (1988). *Angewandtes Problemlöseverhalten—Denken und Handeln in komplexen Zusammenhängen.* Berlin: Springer.

Stowasser, R., & Mohry, B. (1977). Rekursive Verfahren aus der Mathematikgeschichte für den Unterricht. *Der Mathematikunterricht, 23*(1), 5-41.

Zimmermann, B. (1983). Problemlösen als eine Leitidee für den Mathematikunterricht—Ein Bericht über neuere amerikanische Beiträge. *Der Mathematikunterricht, 29*(3), 5-45.